Distributed Rate-Constrained LCMV Beamforming

Jie Zhang, Andreas I. Koutrouvelis, Richard Heusdens, and Richard C. Hendriks

***Abstract*—In this letter, we propose a decentralized framework for rate-distributed linearly constrained minimum variance (LCMV) beamforming in wireless acoustic sensor networks. To save the energy usage within the network, we propose to minimize the transmission cost and put a constraint on the noise reduction performance. Subsequently, we decentralize the obtained LCMV filter structure by exploiting an imposed block diagonal form of the noise correlation matrix. As a result, the beamformer weights are calculated in a decentralized fashion and each node can determine its quantization rate locally. Finally, numerical results validate the proposed method.**

***Index Terms*—Rate allocation, LCMV, noise reduction, energy usage, distributed beamforming, acoustic sensor networks.**

I. Introduction

RECENTLY, several beamforming algorithms for wireless acoustic sensor networks (WASNs) have been proposed [1]–[9]. The calculations are done either in a centralized way [1]–[4] or in a distributed way [5]–[9]. In the centralized case, all the sensor nodes need to transmit their measurements to a fusion center (FC), and the FC performs all computations. There are several limitations on the centralized approach. First, the amount of data that needs to be sent and saved in the FC scales up with the network size. Moreover, with an FC, all operations are performed in a single node, which, in case of disconnection from the network, will cause full collapse of the system. In contrast, the decentralized implementation distributes calculations over the nodes in the WASN, which could overcome the limitations of the centralized approaches.

In WASNs, usually the sensors are battery-powered with a limited energy budget. To reduce the energy consumption of beamforming algorithms, one could apply sensor selection [10]–[12] or rate allocation [13]–[16] to reduce the amount of transmitted information. Rate allocation is more general than sensor selection, as it allows for multiple decisions on the status of sensors. However, sensor selection and rate allocation methods typically work in a centralized fashion, which is, as argued above, undesirable due to scalability and instability issues. In this letter we therefore investigate a decentralized solution for rate-distributed beamforming.

In [17], a distributed linearly constrained minimum variance (LCMV) beamforming method for WASNs was proposed. This method block-diagonalizes the noise/noisy correlation matrix using linear equality constraints, leading to an efficient distributed implementation for the LCMV beamformer. However, this method does not take into account the quantization noise introduced during the communication between the devices. Nor does it take the energy usage due to transmission into account. The rate-distributed LCMV (RD-LCMV) beamformer proposed in [15] is an effective method to reduce the transmission costs over WASNs. It optimally distributes rates to the sensors by minimizing the transmission power under a constraint on the noise reduction performance. However, the RD-LCMV method was derived in a centralized way. This is less efficient with respect to transmission energy if the FC is far away from the communicating sensor.

In this letter our contribution is twofold. First, we solve the rate-allocation problem introduced in [15] for the distributed beamformer proposed in [17]. As the beamformer output highly depends on the quantization noise, we allocate the rates between the devices such that the distributed LCMV beamformer in [17] guarantees a pre-defined performance. Secondly, we propose a distributed solution to the RD-LCMV problem introduced in [15]. Experiments in a simulated WASN validate the proposed decentralized method, i.e., the expected noise reduction performance is achieved with a saving of transmission costs compared to the centralized implementation.

II. Fundamentals

A. Signal Model

We consider a connected WASN consisting of K nodes, where each node $k \in \mathcal{K} = \{1, \ldots, K\}$, with $\mathcal{K}$ the set of node indices, has $M_k, \forall k$ microphones. In total, we have $M = \sum_{k=1}^{K} M_k$ microphones that acquire the sound field consisting of one target source degraded by acoustic background noise. Let $\mathcal{E}$ denote the set of edges of the network and $\mathcal{N}_k$ the set of neighbouring nodes of node k. If and only if $(k, m) \in \mathcal{E}$, the kth and mth nodes can communicate with each other directly. Let l and ω denote the index of time frame and angular frequency, respectively. In the short-term Fourier transform (STFT) domain, the noisy STFT coefficient at the κth microphone, say $Y_\kappa(\omega, l), \forall \kappa$, is given by

$$Y_\kappa(\omega, l) = X_\kappa(\omega, l) + N_\kappa(\omega, l), \quad (1)$$

where $X_\kappa(\omega, l) = a_\kappa(\omega) S(\omega, l)$ with $a_\kappa(\omega)$ the acoustic transfer function (ATF) of the target signal with respect to the κth microphone and $S(\omega, l)$ the STFT coefficient of the target source signal at the source location. In reverberant environments, the ATF consists of early reverberation (typically the first 50 ms) and late reverberation components [18], [19]. Only the early reflections of the target source are beneficial for improving the speech intelligibility [19]. Therefore, in (1), the total noise $N_\kappa(\omega, l)$ received by microphone κ is given by

$$N_\kappa(\omega, l) = Z_\kappa(\omega, l) + U_\kappa(\omega, l), \quad (2)$$

Manuscript received January 22, 2019; revised March 9, 2019; accepted March 11, 2019. Date of publication March 15, 2019; date of current version March 27, 2019. This work was supported by the China Scholarship Council under Grant 201506010331. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Rosângela Fernandes Coelho. *(Corresponding author: Jie Zhang.)*

The authors are with the Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: j.zhang-7@tudelft.nl; a.koutrouvelis@tudelft.nl; r.heusdens@tudelft.nl; r.c.hendriks@tudelft.nl).

Digital Object Identifier 10.1109/LSP.2019.2905161

where $Z_\kappa(\omega, l)$ denotes the correlated noise components including the early reflections of all interfering sources, and $U_k(\omega, l)$ the remaining noise components including the late reverberation from all sources and the sensor noise. For notational brevity, the frequency variable ω and the frame index l will be omitted now onwards. Using vector notation, the M channel signals are stacked in a vector $\mathbf{y} = [Y_1, \ldots, Y_M]^T \in \mathbb{C}^M$. Similarly, we define M-dimensional vectors $\mathbf{x}, \mathbf{n}, \mathbf{z}, \mathbf{u}, \mathbf{a}$ for the clean speech component, the total noise, the correlated noise, remaining noise and ATF, respectively, such that the signal model in (1) can compactly be written as

$$\mathbf{y} = \mathbf{x} + \mathbf{n} = \mathbf{x} + \mathbf{z} + \mathbf{u}, \tag{3}$$

where $\mathbf{x} = \mathbf{a}S$. To focus on the concept of rate-distributed noise reduction, we assume in this letter that the ATFs of all sources are known. In a centralized setting, the RTF can be estimated using covariance substraction or covariance whitening method [20]. In the distributed setting this can be estimated using [21]–[24]. Further, we assume that all sources are mutually uncorrelated, and the early reflections and late reverberation are also mutually uncorrelated (which is strictly speaking true under the assumption that the STFT coefficients S across time are uncorrelated), such that the second-order statistics (SOS) of the noise components can be written as

$$\mathbf{R_n} = \mathbb{E}\left[|\mathbf{n}|^2\right] = \mathbf{R_z} + \mathbf{R_u}, \tag{4}$$

where $\mathbb{E}\{\cdot\}$ denotes the statistical expectation operation.

B. Centralized LCMV Beamforming

The LCMV beamformer [25]–[28] is widely used in array processing. The filter coefficients are designed to minimize the output noise power subject to a set of linear constraints,

$$\mathbf{w}_{\text{LCMV}} = \arg\min_{\mathbf{w}} \ \mathbf{w}^H\mathbf{R_n}\mathbf{w}, \quad \text{s.t.} \quad \mathbf{\Lambda}^H\mathbf{w} = \mathbf{f}. \tag{5}$$

The closed-form solution to (5) is given by [25]–[28]

$$\mathbf{w}_{\text{LCMV}} = \mathbf{R_n}^{-1}\mathbf{\Lambda}\left(\mathbf{\Lambda}^H\mathbf{R_n}^{-1}\mathbf{\Lambda}\right)^{-1}\mathbf{f}. \tag{6}$$

Notably, the linear constraints in (5) can be used to preserve target sources, eliminate interfering sources [25]–[28], or preserve the spatial cues of the sound field [16], [29], [30].

In general, the microphones within a single node are spatially close, while the microphones at different nodes in a WASN are typically more distant. In [17], it was argued that the late reverberation is highly correlated in the first case, while much less correlated in the latter case. Hence, it was suggested that the SOS $\mathbf{R_u}$ can be approximated by a *block-diagonal* matrix where each block corresponds to the SOS of the late reverberation of one node only and the microphone self-noise. By properly using the constraints in the LCMV framework to cancel the early components contained in $\mathbf{z}$ and leveraging the block-diagonal structure of the SOS, the LCMV beamforming problem in (5) can be implemented in a distributed fashion. Hence, as in [17], in this letter we specify $\mathbf{f} = [1, 0, \cdots, 0]^T \in \mathbb{C}^{r+1}$ (r is the number of interferers), and $\mathbf{\Lambda} = [\mathbf{a}, \mathbf{b}_1, \cdots, \mathbf{b}_r] \in \mathbb{C}^{M\times(r+1)}$ consisting of ATF vectors with $\mathbf{b}_j, \forall j$ the ATF of the jth interfering source. Clearly, with such a set of linear constraints $\mathbf{\Lambda}^H\mathbf{w} = \mathbf{f}$ and given enough degrees-of-freedom, the power of the target source is preserved and the power of the correlated sources can entirely be suppressed. As a result, the output noise power after LCMV beamforming can be shown to be given by [28]

$$\mathbb{E}\left[|\mathbf{w}^H\mathbf{n}|^2\right] = \mathbb{E}\left[|\mathbf{w}^H\mathbf{u}|^2\right] = \mathbf{w}^H\mathbf{R_u}\mathbf{w}, \tag{7}$$

due to the fact that $\mathbf{b}_j^H\mathbf{w} = 0, \forall j$. That is, any decrease in the objective function of (5) is caused by reducing the uncorrelated noise components. As a result, the matrix $\mathbf{R_n}$ can be replaced by $\mathbf{R_u}$. In the sequel, we will use the block-diagonal approximation of $\mathbf{R_u}$ for the design of algorithms.

III. Distributed LCMV Beamforming With Quantization Noise

Given the block-diagonal matrix $\mathbf{R_u}$, by using (7) and the constraints to null the early components contained in $\mathbf{z}$, the centralized LCMV beamforming problem in (5) can be written in the following node separable form:

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \sum_{k=1}^{K} \mathbf{w}_k^H\mathbf{R}_{\mathbf{u},k}\mathbf{w}_k, \text{ s.t. } \sum_{k=1}^{K}\mathbf{\Lambda}_k^H\mathbf{w}_k = \mathbf{f}, \tag{8}$$

where $\mathbf{w}_k \in \mathbb{C}^{M_k}$, $\mathbf{\Lambda}_k \in \mathbb{C}^{M_k\times(r+1)}$ and $\mathbf{R}_{\mathbf{u},k} = \mathbb{E}[\mathbf{u}_k\mathbf{u}_k^H] \in \mathbb{C}^{M_k\times M_k}$ with $\mathbf{u}_k \in \mathbb{C}^{M_k}$ denote the elements of $\mathbf{w}$, the rows of $\mathbf{\Lambda}$ and the kth block of the matrix $\mathbf{R_u}$, respectively. The subscript k is used to indicate the components associated with node k. Considering the real-valued Lagrangian function of (8), we can obtain the optimal local LCMV filter, given by [17]

$$\mathbf{w}_k^* = \mathbf{R}_{\mathbf{u},k}^{-1}\mathbf{\Lambda}_k\boldsymbol{\mu}^*, \tag{9}$$

where $\boldsymbol{\mu}^* \in \mathbb{C}^{r+1}$ is a vector with Lagrangian multipliers. Clearly, the optimal local LCMV filter $\mathbf{w}_k^*$ depends on the global optimal dual variables $\boldsymbol{\mu}^*$. To determine $\boldsymbol{\mu}^*$, one can consider the dual optimization problem of (8), given by

$$\boldsymbol{\mu}^* = \arg\max_{\boldsymbol{\mu}} -\sum_{k=1}^{K}\boldsymbol{\mu}^H\mathbf{\Lambda}_k^H\mathbf{R}_{\mathbf{u},k}^{-1}\mathbf{\Lambda}_k\boldsymbol{\mu} + 2\Re\left(\boldsymbol{\mu}^H\mathbf{f}\right), \tag{10}$$

where $\Re(\cdot)$ returns the real part. For notational simplicity, we define $\mathbf{G}_k = \mathbf{\Lambda}_k^H\mathbf{R}_{\mathbf{u},k}^{-1}\mathbf{\Lambda}_k, \forall k$. To optimize (10) in a distributed fashion, we introduce $\boldsymbol{\mu}_k, \forall k$ to denote the local version of $\boldsymbol{\mu}$ at each node. With this, (10) is equivalent to

$$\min_{\boldsymbol{\mu}_k} \sum_{k=1}^{K}\left(\boldsymbol{\mu}_k^H\mathbf{G}_k\boldsymbol{\mu}_k - \frac{2}{K}\Re\left(\boldsymbol{\mu}_k^H\mathbf{f}\right)\right) \text{ s.t. } \boldsymbol{\mu}_k = \boldsymbol{\mu}_m,$$

for all $(k, m) \in \mathcal{E}$. The resulting problem can be solved using randomized gossip [31], ADMM [32] or PDMM [33]. For instance, as shown in [17], the PDMM update procedure for the $(i+1)$th iteration can be summarized as

$$\begin{aligned}\boldsymbol{\mu}_k^{(i+1)} = {} & (\mathbf{G}_k + \rho|\mathcal{N}_k|\mathbf{I})^{-1} \\ & \times\left[\sum_{m\in\mathcal{N}_k}\left(\frac{k-m}{|k-m|}\boldsymbol{\gamma}_{m|k}^{(i)} + \rho\boldsymbol{\mu}_m^{(i)}\right) + \frac{\mathbf{f}}{K}\right],\end{aligned} \tag{11a}$$

$$\boldsymbol{\gamma}_{k|m}^{(i+1)} = \boldsymbol{\gamma}_{m|k}^{(i)} - \rho\frac{k-m}{|k-m|}\left(\boldsymbol{\mu}_k^{(i+1)} - \boldsymbol{\mu}_m^{(i)}\right), \tag{11b}$$

where $\boldsymbol{\gamma}_{k|m}$ and $\boldsymbol{\gamma}_{m|k}$ are the direct-edge variables computed at nodes k and m, respectively, associated with the edge $(k, m) \in \mathcal{E}$, $\mathbf{I}$ denotes the identity matrix, and ρ is a positive step size. Note that in (11), by substituting the update equation for $\boldsymbol{\gamma}_{m|k}^{(i)}$, we can get rid of transmitting the edge variables. As such, updating the edge variables can be performed by broadingcasting $\boldsymbol{\mu}_k^{(i)}$. The iterative procedure can be terminated until $|\boldsymbol{\mu}_k^{(i)} - \boldsymbol{\mu}_m^{(i)}| < \epsilon$ where ϵ is a small positive number.

In [34], [35], the convergence of PDMM was shown in the presence of quantization noise. Due to quantization, the dual variables exchanged among nodes are noisy, i.e., $\hat{\boldsymbol{\mu}}_k^{(i)} = \boldsymbol{\mu}_k^{(i)} + \tilde{\boldsymbol{\mu}}_k^{(i)}$, where $\tilde{\boldsymbol{\mu}}_k^{(i)}$ denotes the quantization noise which is assumed to be zero-mean.[1] Using the above PDMM update equations, the LCMV filter from (9) in iteration i is given by

$$\hat{\mathbf{w}}_k^{(i)} = \mathbf{w}_k^{(i)} + \tilde{\mathbf{w}}_k^{(i)} = \mathbf{R}_{\mathbf{u},k}^{-1}\boldsymbol{\Lambda}_k\left(\boldsymbol{\mu}_k^{(i)} + \tilde{\boldsymbol{\mu}}_k^{(i)}\right), \tag{12}$$

where $\tilde{\mathbf{w}}_k^{(i)} = \mathbf{R}_{\mathbf{u},k}^{-1}\boldsymbol{\Lambda}_k\tilde{\boldsymbol{\mu}}_k^{(i)}$ is the error caused by quantization. After the local filters are obtained, calculating the beamformer output reduces to an average consensus problem as

$$\min_X \sum_{k=1}^{K}(X_k - \hat{\mathbf{w}}_k^H\mathbf{y}_k)^2 \text{ s.t. } X_k = X_m, \forall(k,m) \in \mathcal{E}. \tag{13}$$

The PDMM update equations for (13) can be found in [17]. Note that for stationary signals, the update procedure in (11) is time-invariant, while (13) is always both time and frequency dependent. To reduce the communication costs, we will next derive how to find the optimal quantization rate distribution for iteratively calculating the local filters and beamforming.

IV. Proposed Distributed Rate Allocation

Let the transmission power from node k to a neighboring node m for a single time-frequency bin be $d_k^2 V_{km}(4^{b_k} - 1)$, where $0 \le b_k \le b_0, \forall k$ denotes the integer rate that is used by the node k, and d_k and V_{km} denote the transmission range and the channel noise power spectral density (PSD) between node k and node m, respectively [36]–[38]. Assuming that in each iteration we randomly (e.g., at a probability of $\frac{1}{K}$) pick one node of the WASN that broadcasts information to all of its neighboring nodes, such that the expected transmission power per iteration can be given by

$$g(\mathbf{b}) = \frac{1}{K}\sum_{k=1}^{K} d_k^2 V_k(4^{b_k} - 1), \tag{14}$$

where V_k is the mean value of $V_{km}, m \in \mathcal{N}_k$. Assuming that I iterations are used for calculating the filters through (11) and J iterations for beamforming in (13), respectively, the original RD-LCMV problem in [15] can be reformulated as

$$\min_{\mathbf{b}} \; g(\mathbf{b}) \text{ s.t. } \sum_{k=1}^{K}\left(\mathbb{E}\left[\left|\hat{\mathbf{w}}_k^{(I)H}\mathbf{u}_k\right|^2\right] + \mathbb{E}\left[\zeta_{X_k}^{(J)}\right]\right) \le \frac{\beta}{\alpha}, \tag{P1}$$

where $\alpha \in (0,1]$ is the parameter to control the expected performance, $\mathbb{E}[\zeta_{X_k}^{(J)}]$ denotes the primal mean-squared error (MSE) caused by quantizing X_k in calculating the beamformer output, i.e., $\zeta_{X_k}^{(J)} = |X_k - Q_{b_k}^{(J)}(X_k)|^2$ with $Q_{b_k}^{(J)}(X_k)$ denoting the quantized X_k using b_k bits. Further, the filter $\hat{\mathbf{w}}_k^{(I)}$ was given in (12), and $\beta = \sum_{k=1}^{K}\mathbb{E}[|\mathbf{w}_k^{(I)H}\mathbf{u}_k|^2]$ denotes the minimum output noise power (i.e., without quantization noise). In (P1), the term $\mathbb{E}[|\hat{\mathbf{w}}_k^{(I)H}\mathbf{u}_k|^2]$ denotes the residual acoustic noise and the residual noise of the beamformer due to quantizing $\boldsymbol{\mu}_k$. Note that $\zeta_{X_k}^{(J)}$ depends on the number of iterations and the topology of the network. Since the beamforming is performed iteratively with quantization, the quantization noise $\zeta_{X_k}^{(J)}$ will accumulate at each iteration. However, in [34], it was shown that in case of quantization with sufficiently small fixed cell width (e.g., uniform quantization), the error accumulates but the growth is so slow that it can be considered constant over the iteration range of interest. That is, the primal MSE $\mathbb{E}[\zeta_{X_k}^{(J)}]$ can be approximated by

$$\mathbb{E}\left[\zeta_{X_k}^{(J)}\right] \approx C\sigma_k^2, \forall k, \tag{15}$$

where σ_k^2 denotes the noise variance depending on the bit rate and the quantization range, and C is a constant which only depends on the topology of the network and is $\mathcal{O}(K)$.

The noise power at node k in (P1) can be calculated by

$$\begin{aligned}
\mathbb{E}\left[\left|\hat{\mathbf{w}}_k^{(I)H}\mathbf{u}_k\right|^2\right] &\overset{\text{(a)}}{=} \mathbb{E}\left[\left(\mathbf{w}_k^{(I)} + \tilde{\mathbf{w}}_k^{(I)}\right)^H \mathbf{u}_k\mathbf{u}_k^H\left(\mathbf{w}_k^{(I)} + \tilde{\mathbf{w}}_k^{(I)}\right)\right] \\
&\overset{\text{(b)}}{=} \mathbb{E}\left[\mathbf{w}_k^{(I)H}\mathbf{u}_k\mathbf{u}_k^H\mathbf{w}_k^{(I)}\right] + 2\mathbb{E}\left[\Re\left(\mathbf{w}_k^{(I)H}\mathbf{u}_k\mathbf{u}_k^H\tilde{\mathbf{w}}_k^{(I)}\right)\right] \\
&\quad + \mathbb{E}\left[\tilde{\mathbf{w}}_k^{(I)H}\mathbf{u}_k\mathbf{u}_k^H\tilde{\mathbf{w}}_k^{(I)}\right],
\end{aligned}$$

where we note that $\sum_{k=1}^{K}\mathbb{E}[\mathbf{w}_k^{(I)H}\mathbf{u}_k\mathbf{u}_k^H\mathbf{w}_k^{(I)}] = \beta$.

Proposition 1: If the quantization noise $\tilde{\boldsymbol{\mu}}_k^{(I)}$ and the acoustic noise $\mathbf{u}_k$ are independent, we have $\mathbb{E}[\Re(\mathbf{w}_k^{(I)H}\mathbf{u}_k\mathbf{u}_k^H\tilde{\mathbf{w}}_k^{(I)})] = 0$ and $\mathbb{E}[\tilde{\mathbf{w}}_k^{(I)H}\mathbf{u}_k\mathbf{u}_k^H\tilde{\mathbf{w}}_k^{(I)}] = \mathrm{Tr}(\mathbf{G}_k\mathbf{R}_{\tilde{\boldsymbol{\mu}}_k})$ where $\mathbf{R}_{\tilde{\boldsymbol{\mu}}_k} = \mathbb{E}[\tilde{\boldsymbol{\mu}}_k^{(I)}\tilde{\boldsymbol{\mu}}_k^{(I)H}]$ and $\mathrm{Tr}(\cdot)$ returns the trace of a matrix.

Proof: The proof follows from the observation that $\mathbb{E}(AB) = \mathbb{E}(A)\mathbb{E}(B)$ if A and B are independent (and $\mathbb{E}(\mathbf{a}^H\mathbf{b}\mathbf{b}^H\mathbf{a}) = \mathbb{E}(\mathrm{Tr}(\mathbf{b}\mathbf{b}^H\mathbf{a}\mathbf{a}^H)) = \mathrm{Tr}(\mathbb{E}(\mathbf{a}\mathbf{a}^H)\mathbb{E}(\mathbf{b}\mathbf{b}^H))$ if $\mathbf{a}$ and $\mathbf{b}$ are independent vectors). ∎

To this end, we can see that

$$\sum_{k=1}^{K}\mathbb{E}\left[\left|\hat{\mathbf{w}}_k^{(I)H}\mathbf{u}_k\right|^2\right] = \beta + \sum_{k=1}^{K}\mathrm{Tr}\left(\mathbf{G}_k\mathbf{R}_{\tilde{\boldsymbol{\mu}}_k}\right). \tag{16}$$

Further, we use fixed-rate uniform quantizers for all iterations to quantize the dual variables, such that the SOS of the quantization noise $\tilde{\boldsymbol{\mu}}_k^{(I)}$ can be given by [14], [15], [39]

$$\mathbf{R}_{\tilde{\boldsymbol{\mu}}_k} = \mathbb{E}\left[\tilde{\boldsymbol{\mu}}_k^{(i)}\tilde{\boldsymbol{\mu}}_k^{(i)H}\right] = \frac{1}{12} \times \frac{\mathcal{A}^2}{4^{b_k}}\mathbf{I}_{r+1}, \forall i, \tag{17}$$

where $\mathcal{A} = \max|\boldsymbol{\mu}^*|$. Similarly, we have $\sigma_k^2 = \frac{1}{12} \times \frac{\mathcal{B}_k^2}{4^{b_k}}$ with $\mathcal{B}_k = \max|\mathbf{w}_k^H\mathbf{y}_k|$. Using a variable change $1 \le t_k = 4^{b_k} \le 4^{b_0}, \forall k$ and the property in (15), (P1) can be reformulated as

$$\min_{\mathbf{t}} \; g(\mathbf{b}) \text{ s.t. } \sum_{k=1}^{K}\left[\mathrm{Tr}\left(\mathbf{G}_k\right)\mathcal{A}^2 + \mathcal{B}_k^2 C\right]/t_k \le \delta, \tag{P2}$$

[1]This assumption holds when subtractive dithering based uniform quantization is used. The dither signal, which is known at the receiver side, and the quantization noise are i.i.d. processes.

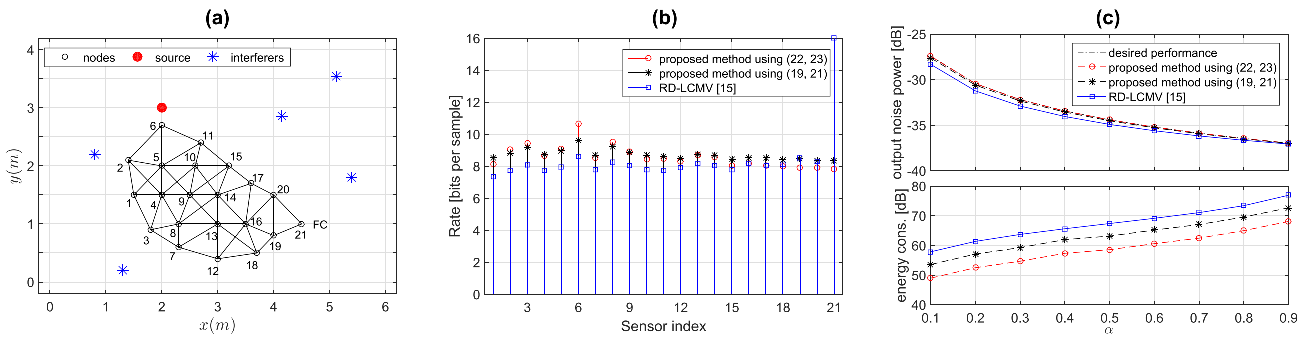

Fig. 1. (a) Experimental setup, where the last node is assumed to be the FC for the centralized RD-LCMV method [15], (b) rate distribution for one frequency bin with $\alpha = 0.8$ and $b_0 = 16$ bits per sample, (c) output noise power and transmission energy in terms of α.

where $\delta = 12(\frac{\beta}{\alpha} - \beta)$. By solving the KKT condition $\frac{\partial \mathcal{L}(\mathbf{t},\lambda)}{\partial t_k} = 0$, the optimal solution to (P2) can be found as

$$t_k^* = \sqrt{\lambda(\mathcal{A}^2 \mathrm{Tr}(\mathbf{G}_k) + \mathcal{B}_k^2 C)/d_k^2 V_k}, \tag{18}$$

which only depends on the Lagrange multiplier λ. To determine λ, one can consider the dual problem of (P2). Substituting (18) into (P2), we obtain the dual problem as

$$\min_{\lambda} \sum_{k=1}^{K} \left(\frac{\delta}{K}\lambda - 2\sqrt{\lambda(\mathcal{A}^2 \mathrm{Tr}(\mathbf{G}_k) + \mathcal{B}_k^2 C) d_k^2 V_k} + d_k^2 V_k \right), \tag{19}$$

which is quadratic in $\sqrt{\lambda}$ and the constraint on δ is partitioned into K equal parts. As a result, we can see that the optimal global multiplier is given by

$$\lambda^* = \frac{1}{K^2} \left(\sum_{k=1}^{K} \sqrt{\lambda_k} \right)^2, \tag{20}$$

where the local λ_k is defined by

$$\lambda_k = K^2 (\mathcal{A}^2 \mathrm{Tr}(\mathbf{G}_k) + \mathcal{B}_k^2 C) d_k^2 V_k / \delta^2, \forall k. \tag{21}$$

Clearly, determining λ^* turns into an averaging problem, since λ_k can be computed separately at each node. Then, we can use PDMM to calculate the average consensus of $\sqrt{\lambda_k}$ that is required by (20). This requires a large amount of information exchange. To avoid this, we can consider using the locally optimal λ_k from (21) only, instead of the globally optimal λ^*. Substituting (21) into (18), we obtain the rate distribution as

$$t_k = K(\mathcal{A}^2 \mathrm{Tr}(\mathbf{G}_k) + \mathcal{B}_k^2 C)/\delta, \tag{22}$$

which reveals that by using local λ_k, the rate can be determined locally without any information exchange and it only depends on the noise power. However, this might affect the global optimality of the rate distribution, which will be studied experimentally. Notably, the final rates should be resolved by $b_k = \log_4 t_k, \forall k$ and randomized rounding as in [15].

V. Numerical Results

Fig. 1(a) shows a simulated WASN in a 2D room with dimensions (6×4) m. We consider $K = 21$ nodes and each node has $M_k = 3, \forall k$ microphones. We set $\rho = 0.5$ and $C = 21$. One target source is located at (2, 3) m. Five noise sources are randomly placed around the WASN. The duration of all sources is 10 minutes. All sources originate from the TIMIT database [40]. The sensor noise is modeled as white Gaussian noise at an SNR of 50 dB. The sampling frequency is 16 kHz. A square-root-Hann window of 50 ms for framing with 50% overlap is applied to the signals. The ATFs are generated using [41] with reverberation time $T_{60} = 200$ ms. The 21st node is assumed to be the FC for the centralized RD-LCMV method [15], i.e., all other nodes are only connected to this FC.

When we calculate the dual variable $\boldsymbol{\mu}$ using PDMM from (11), the warm-start procedure proposed in [17] is employed to achieve an acceptable precision of PDMM within a finite number of iterations. Fig. 1(b) shows a rate-distribution example of the proposed method and the centralized method [15] for $\alpha = 0.8$. For the proposed method, the nodes that have higher SNR are allocated with higher rate, e.g., node 6. For the centralized method [15], the nodes that are closer to the FC are allocated with higher rate. In addition, we show the output noise power and transmission cost averaged over frequencies in terms of α in Fig. 1(c). The energy of the RD-LCMV method is used for transmitting the raw audio realizations. For the proposed method, if we use the local λ_k in (21) to determine the rate distribution, the energy is only used for transmitting the dual variable $\boldsymbol{\mu}$ and calculating the beamformer output; if the rate distribution is computed using (18) with the global λ^* from (20), some extra energy needs to be spent for calculating λ^*. Clearly, both the centralized method and the proposed decentralized method satisfy the desired noise reduction performance, while the proposed method using (21)-(22) consumes less energy, since each sensor node only needs to communicate with the neighboring nodes, instead of with the remote FC. This reveals that using the local λ_k is effective for the energy usage versus performance trade-off in spite of scarifying rate optimality. Note that in general a global optimization problem cannot be approached by optimizing local sub-problems separately. We considered optimizing the local problems in this letter, as the simulation results show that it gives a better energy usage versus performance trade-off.

VI. Conclusions

In this letter, we solved the rate-distributed LCMV beamforming problem in [15] in a fully distributed fashion. The quantization rates were determined locally without any information exchange. Numerical results show the superiority of the proposed method in energy usage. More importantly, the decentralized implementation is more robust against the network variation compared to the centralized method.

REFERENCES

[1] I. Himawan, I. McCowan, and S. Sridharan,“ Clustered blind beamforming from ad-hoc microphone arrays,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 4, pp. 661–676, May 2011.

[2] V. M. Tavakoli, J. R. Jensen, R. Heusdens, J. Benesty, and M. G. Christensen, “Ad hoc microphone array beamforming using the primal-dual method of multipliers,” in *Proc. Eur. Signal Process. Conf.*, 2016, pp. 1088–1092.

[3] S. Gannot, E. Vincent, S. Markovich-Golan, and A. Ozerov, “A consolidated perspective on multimicrophone speech enhancement and source separation,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 25, no. 4, pp. 692–730, Apr. 2017.

[4] V. M. Tavakoli, J. R. Jensen, M. G. Christensen, and J. Benesty, “A framework for speech enhancement with ad hoc microphone arrays,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 6, pp. 1038–1051, Jun. 2016.

[5] A. Bertrand and M. Moonen, “Distributed node-specific LCMV beamforming in wireless sensor networks,” *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 233–246, Jan. 2012.

[6] R. Heusdens, G. Zhang, R. C. Hendriks, Y. Zeng, and W. B. Kleijn, “Distributed MVDR beamforming for (wireless) microphone networks using message passing,” in *Proc. Int. Workshop Acoustic Echo, Noise Control*, 2012, pp. 1–4.

[7] S. Markovich-Golan, S. Gannot, and I. Cohen, “Distributed GSC beamforming using the relative transfer function,” in *Proc. Eur. Signal Process. Conf.*, Aug. 2012, pp. 1274–1278.

[8] Y. Zeng and R. C. Hendriks, “Distributed delay and sum beamformer for speech enhancement via randomized gossip,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 1, pp. 260–273, Jan. 2014.

[9] T. Sherson, W. B. Kleijn, and R. Heusdens, “A distributed algorithm for robust LCMV beamforming.,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2016, pp. 101–105.

[10] S. P. Chepuri and G. Leus, “Sparsity-promoting sensor selection for non-linear measurement models,” *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 684–698, Feb. 2015.

[11] J. Zhang, S. P. Chepuri, R. C. Hendriks, and R. Heusdens, “Microphone subset selection for MVDR beamformer based noise reduction,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 3, pp. 550–563, Mar. 2018.

[12] S. Joshi and S. Boyd, “Sensor selection via convex optimization,” *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 451–462, Feb. 2009.

[13] O. Roy and M. Vetterli, “Rate-constrained collaborative noise reduction for wireless hearing aids,” *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 645–657, Feb. 2009.

[14] F. de la Hucha Arce, F. Rosas, M. Moonen, M. Verhelst, and A. Bertrand, “Generalized signal utility for LMMSE signal estimation with application to greedy quantization in wireless sensor networks,” *IEEE Signal Process. Lett.*, vol. 23, no. 9, pp. 1202–1206, Sep. 2016.

[15] J. Zhang, R. Heusdens, and R. C. Hendriks, “Rate-distributed spatial filtering based noise reduction in wireless acoustic sensor networks,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 11, pp. 2015–2026, Nov. 2018.

[16] J. Zhang, R. Heusdens, and R. C. Hendriks, “Rate-distributed binaural LCMV beamforming for assistive hearing in wireless acoustic sensor networks,” in *Proc. IEEE 10th Sensor Array Multichannel Signal Process. Workshop*, 2018, pp. 460–464.

[17] A. I. Koutrouvelis, T. W. Sherson, R. Heusdens, and R. C. Hendriks, “A low-cost robust distributed linearly constrained beamformer for wireless acoustic sensor networks with arbitrary topology,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 8, pp. 1434–1448, Aug. 2018.

[18] J. L. Flanagan, A. C. Surendran, and E. Jan, “Spatially selective sound capture for speech and audio processing,” *Speech Commun.*, vol. 13, nos. 1–2, pp. 207–222, 1993.

[19] J. Bradley, H. Sato, and M. Picard, “On the importance of early reflections for speech in rooms,” *J. Acoust. Soc. Amer.*, vol. 113, no. 6, pp. 3233–3244, 2003.

[20] S. Markovich-Golan, S. Gannot, and W. Kellermann, “Performance analysis of the covariance-whitening and the covariance-subtraction methods for estimating the relative transfer function,” in *Proc. Eur. Signal Process. Conf.*, 2018, pp. 2513–2517.

[21] S. Markovich Golan, S. Gannot, and I. Cohen, “A reduced bandwidth binaural MVDR beamformer,” in *Proc. Int. Workshop Acoustic Echo, Noise Control*, 2010.

[22] S. Markovich-Golan, A. Bertrand, M. Moonen, and S. Gannot, “Optimal distributed minimum-variance beamforming approaches for speech enhancement in wireless acoustic sensor networks,” *Signal Process.*, vol. 107, pp. 4–20, 2015.

[23] A. Bertrand and M. Moonen, “Distributed adaptive generalized eigenvector estimation of a sensor signal covariance matrix pair in a fully connected sensor network,” *Signal Process.*, vol. 106, pp. 209–214, 2015.

[24] A. Hassani, J. Plata-Chaves, M. H. Bahari, M. Moonen, and A. Bertrand, “Multi-task wireless sensor network for joint distributed node-specific signal enhancement, LCMV beamforming and DOA estimation,” *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 3, pp. 518–533, Apr. 2017.

[25] O. L. Frost III, “An algorithm for linearly constrained adaptive array processing,” *Proc. IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1972.

[26] B. D. Van Veen and K. M. Buckley, “Beamforming: A versatile approach to spatial filtering,” *IEEE Signal Process. Mag.*, vol. 5, no. 2, pp. 4–24, Apr. 1988.

[27] J. Benesty, S. Makino, and J. Chen, *Speech Enhancement*. Berlin, Germany: Springer Science & Business Media, 2005.

[28] M. Souden, J. Benesty, and S. Affes, “A study of the LCMV and MVDR noise reduction filters,” *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4925–4935, Sep. 2010.

[29] A. I. Koutrouvelis, R. C. Hendriks, R. Heusdens, and J. Jensen, “Relaxed binaural LCMV beamforming,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 25, no. 1, pp. 137–152, Jan. 2017.

[30] E. Hadad, S. Doclo, and S. Gannot, “The binaural LCMV beamformer and its performance analysis,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 3, pp. 543–558, Mar. 2016.

[31] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, “Randomized gossip algorithms,” *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2508–2530, Jun. 2006.

[32] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[33] G. Zhang and R. Heusdens, “Distributed optimization using the primal-dual method of multipliers,” *IEEE Trans. Signal Inf. Process. Netw.*, vol. 4, no. 1, pp. 173–187, Mar. 2018.

[34] D. H. M. Schellekens, T. Sherson, and R. Heusdens, “Quantisation effects in PDMM: A first study for synchronous distributed averaging,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2017, pp. 4237–4241.

[35] J. A. G. Jonkman, T. Sherson, and R. Heusdens, “Quantisation effects in distributed optimisation,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2018, pp. 3649–3653.

[36] S. Shah and B. Beferull-Lozano, “Adaptive quantization for multihop progressive estimation in wireless sensor networks,” in *Proc. Eur. Signal Process. Conf.*, 2013, pp. 1–5.

[37] Y. Huang and Y. Hua, “Multihop progressive decentralized estimation in wireless sensor networks,” *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 1004–1007, Dec. 2007.

[38] Y. Huang and Y. Hua, “Energy planning for progressive estimation in multihop sensor networks,” *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4052–4065, Oct. 2009.

[39] A. Sripad and D. Snyder, “A necessary and sufficient condition for quantization errors to be uniform and white,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-25, no. 5, pp. 442–448, Oct. 1977.

[40] J. S. Garofolo, “DARPA TIMIT acoustic-phonetic speech database,” *Nat. Inst. Standards Technol.*, vol. 15, pp. 29–50, 1988.

[41] E. A. P. Habets, “Room impulse response generator,” Technische Univ. Eindhoven, Eindhoven, The Netherlands, Tech. Rep., 2006, vol. 2, no. 2.4, p. 1.